

Review: Structured-based activity prediction for an enzyme of unknown function by J.C. Hermann *et al.*

Juan Miguel Cejuela <jmcejuela@gmail.com>

Lab Protein Structure and Function Analysis
Lecturers: Dr. Andrea Schafferhans and Dr. Marco Punta

Technische Universität München
Master of Science Informatics

April 25, 2010

Paper Meta Info

Structured-based activity prediction for an enzyme of unknown function

- *authors*: Johannes C. Hermann, Ricardo Marti-Arbona, Alexander A. Fedorov, Elena Fedorov, Steven C. Almo, Brian K. Shoichet, and Frank M. Raushel.
- *publication*: Nature. 2007 August 16; 448(7155): 775-779.
- *url*: <http://www.ncbi.nlm.nih.gov/pubmed/17603473>

Related paper: Predicting substrates by docking high-energy intermediates to enzyme structures, by J.C. Hermann *et al.*

Abstract's Abstract

- bioinformatics inference function difficult and unreliable for proteins without homologues of known function

Abstract's Abstract

- bioinformatics inference function difficult and unreliable for proteins without homologues of known function
- idea: **predict the function** of an enzyme of unknown activity by **docking high-energy intermediate forms** of thousands of potential substrates.

Abstract's Abstract

- bioinformatics inference function difficult and unreliable for proteins without homologues of known function
- idea: **predict the function** of an enzyme of unknown activity by **docking high-energy intermediate forms** of thousands of potential substrates.
- experiment case: enzyme **Tm0936** from Bacteria *Thermotoga maritima*

Structural Complementarity

Structural complementarity is an attractive method when the protein/enzyme has no orthologues with known activity, making inference unreliable.

Structural Complementarity

Structural complementarity is an attractive method when the protein/enzyme has no orthologues with known activity, making inference unreliable.

However, it's difficult.

Structural Complementarity

Structural complementarity is an attractive method when the protein/enzyme has no orthologues with known activity, making inference unreliable.

However, it's difficult.

In addition to problems of sampling and scoring in docking, substrate prediction confronts several additional challenges.

Mainly:

- many possible substrates to consider
- many reactions that enzyme might catalyse
- Enzymes preferably recognize transition states over the ground state structures (the ones in docking databases)

Reduce Search Space

- **Focus on a single class of reactions.** Reduce the number of possible reactions.

Reduce Search Space

- **Focus on a single class of reactions.** Reduce the number of possible reactions.
- Focus on a single source of likely substrates (see referenced related paper)
- Although substrate identification remains challenging, it is at least a **finite problem**

Modelling substrates as high-energy intermediates

Transform ground state structures into structures mimicking the high-energy intermediates that occur along the enzyme reaction coordinate. We refer to these transition-state-like geometries as **high-energy intermediates**

Modelling substrates as high-energy intermediates

Transform ground state structures into structures mimicking the high-energy intermediates that occur along the enzyme reaction coordinate. We refer to these transition-state-like geometries as **high-energy intermediates**

- High-energy intermediates (compared to ground state structures) encode more of the information necessary to specify a fit.

Modelling substrates as high-energy intermediates

Transform ground state structures into structures mimicking the high-energy intermediates that occur along the enzyme reaction coordinate. We refer to these transition-state-like geometries as **high-energy intermediates**

- High-energy intermediates (compared to ground state structures) encode more of the information necessary to specify a fit.
- Among the best complements for the steric and electronic features of the enzyme active site

Enzyme Tm0936

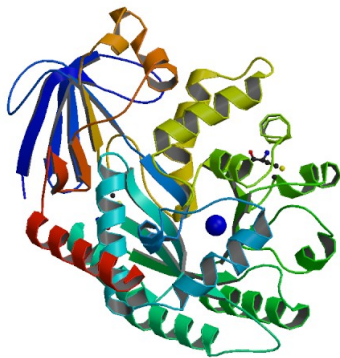


Figure: 1J6P PDB

AHS family

Focus on the reactions catalysed by the amidohydrolase superfamily (AHS), of which Tm0936 is a member.

AHS family

Focus on the reactions catalysed by the amidohydrolase superfamily (AHS), of which Tm0936 is a member.

Big group: 6,000 catalogued members of the AHS which catalyse 30% of reactions in biosynthetic and catabolic pathways.

Enzyme Tm0936 : procedures

All potential groups recognized by AHS enzymes transformed into high-energy intermediates (about 22,500)

Enzyme Tm0936 : procedures

All potential groups recognized by AHS enzymes transformed into high-energy intermediates (about 22,500)

Docked into the structure of Tm0936. Sampled thousands of configurations and conformations of each molecule.

Enzyme Tm0936 : procedures

All potential groups recognized by AHS enzymes transformed into high-energy intermediates (about 22,500)

Docked into the structure of Tm0936. Sampled thousands of configurations and conformations of each molecule.

Best-ranked molecules:

dominated by adenine analogues which appeared to undergo C6-deamination (**tentative function**)

Enzyme Tm0936 : procedures

All potential groups recognized by AHS enzymes transformed into high-energy intermediates (about 22,500)

Docked into the structure of Tm0936. Sampled thousands of configurations and conformations of each molecule.

Best-ranked molecules:

dominated by adenine analogues which appeared to undergo C6-deamination (**tentative function**)

nearly ideal interactions with the active site

Enzyme Tm0936 : procedures

All potential groups recognized by AHS enzymes transformed into high-energy intermediates (about 22,500)

Docked into the structure of Tm0936. Sampled thousands of configurations and conformations of each molecule.

Best-ranked molecules:

dominated by adenine analogues which appeared to undergo C6-deamination (**tentative function**)

nearly ideal interactions with the active site

High-energy intermediate structures typically outperformed docking of the ground state forms of the same molecules

Check now substrates

Best scored high-energy intermediate fits → Potential substrates

Check now substrates

Best scored high-energy intermediate fits → Potential substrates

Among those 5-methylthioadenosine (MTA) and S-adenosylhomocysteine (SAH), which had substantial catalytic rate constants ($10^5 M^{-1} s^{-1}$).

Success

The X-ray crystal structure of the complex between Tm0936 and the product resulting from the deamination of SAH, S-inosylhomocysteine (SIH), was determined, and it corresponded closely to the predicted structure.

Resulting Complex

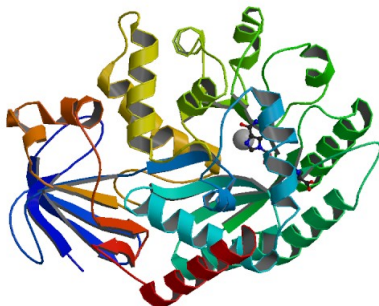


Figure: 2PLM PDB, 2.1 Å resolution by X-ray crystallography

Conclusions for Tm0936 ?

No other enzyme in *T. maritima* has been identified that serves this role



Conclusions for Tm0936 ?

No other enzyme in *T. maritima* has been identified that serves this role



possibility that Tm0936 functions as an adenosine deaminase in *T. maritima*



Conclusions for Tm0936 ?

No other enzyme in *T. maritima* has been identified that serves this role



possibility that Tm0936 functions as an adenosine deaminase in *T. maritima*



Orthologues for Tm0936 may now be classified as MTA/SAH/adenosine deaminases as well (on the basis of evolutionary conservation)

Caveats

Successful function prediction but some caveats:

- Recognition of Tm0936 as an amidohydrolase limited the number of possible reactions to be considered. Not so evident for other cases.
- Restricting to metabolites is not always appropriate
- Tm0936 experienced little conformational change between the apo structure and that of the product complex.

Structured-based docking of high-energy intermediates

But it's unlikely that Tm0936 represents an isolated case.

Whereas conformational change remains a serious challenge, retrospective studies suggest that it is not insurmountable.

Structured-based docking of high-energy intermediates

But it's unlikely that Tm0936 represents an isolated case.

Whereas conformational change remains a serious challenge, retrospective studies suggest that it is not insurmountable.

Structured-based docking of high-energy intermediates should be a useful tool to decrypt the activity of enzymes of unknown function, and will be especially interesting for those targets where bioinformatics inferences is